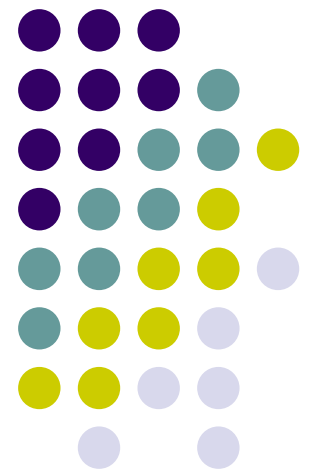
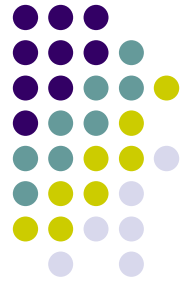


B5 Online MEAP

Paul Stemmer
NAEP State Coordinator

28 Feb 06

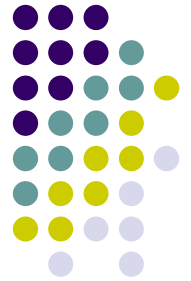




Current NAEP Activities

- Sample was selected last summer
- National Sampling Year
- Verified starting in December
- 75% of Schools finished Testing
- 26 Schools
- Civics, History, (4,8,12)
- Economics (12)

NAEP and MEAP Differences

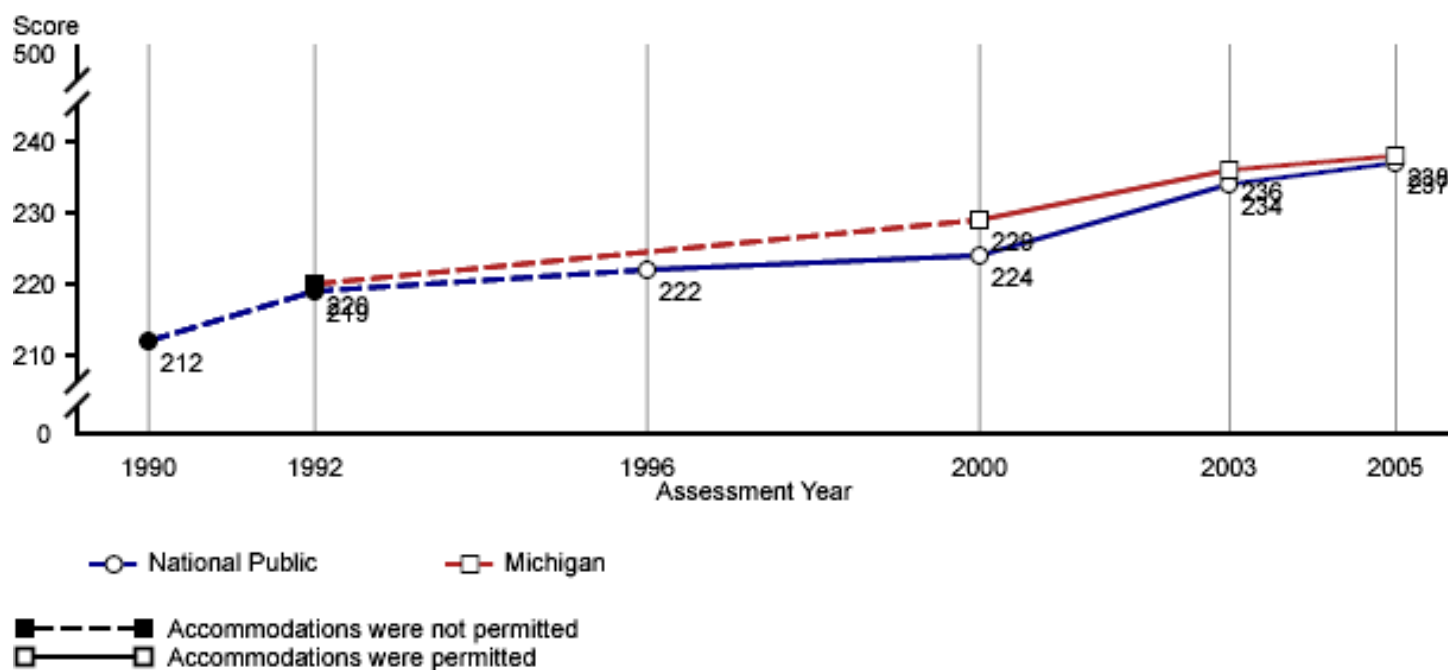


- Sample of districts, schools, students. Uses matrix sample, 90 minutes
- More constructed response
- Test Content Framework – extends some content
- Used for Research and National Policy & Funding Issues
- Every student tested, unlimited time, same questions, matrix future and extended core
- Grade Level Content Expectations & Benchmarks
- Used for accountability

NAEP Grade 4 Mathematics Achievement

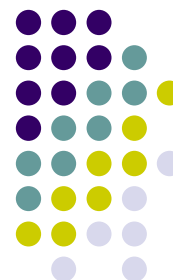


Average scale scores for mathematics, grade 4
All students [TOTAL] = All students
By jurisdiction, 1990, 1992, 1996, 2000, 2003 and 2005

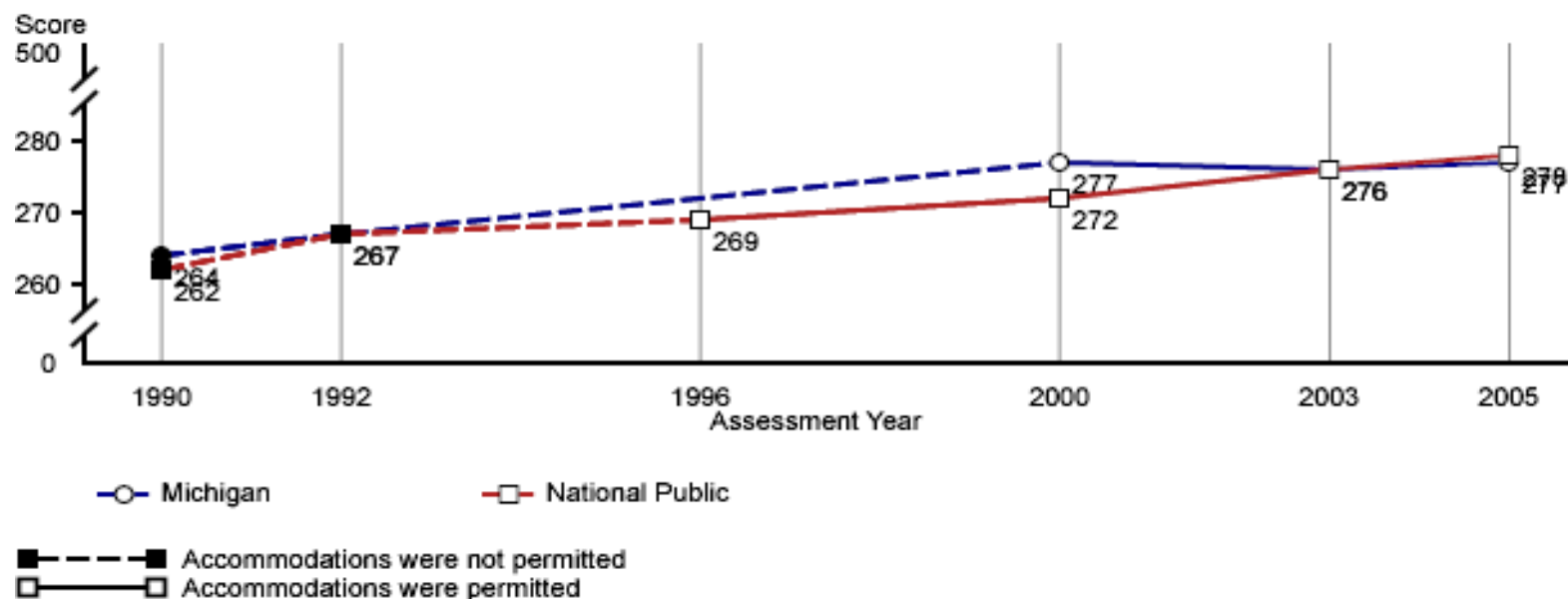


NOTE: The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990, 1992, 1996, 2000, 2003 and 2005 Mathematics Assessments.

NAEP Grade 8 Mathematics Achievement



Average scale scores for mathematics, grade 8
All students [TOTAL] = All students
By jurisdiction, 1990, 1992, 1996, 2000, 2003 and 2005



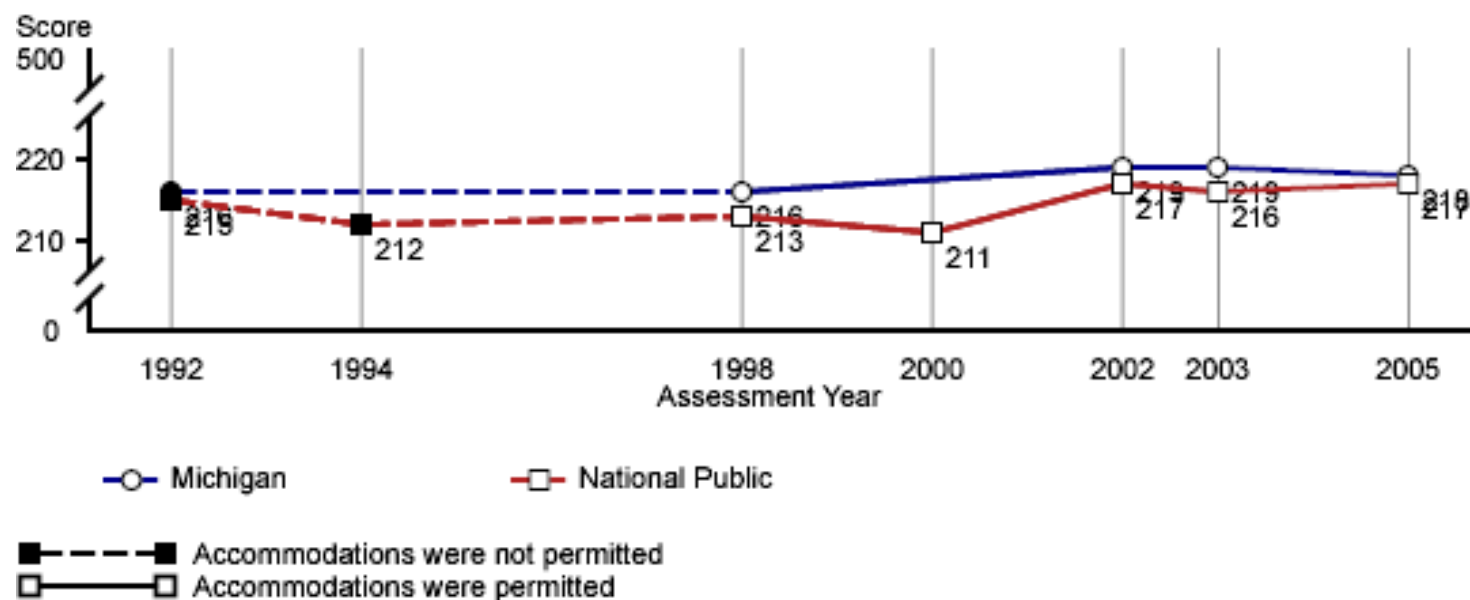
NOTE: The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990, 1992, 1996, 2000, 2003 and 2005 Mathematics Assessments.

NAEP Grade 4 Reading Achievement



Average scale scores for reading, grade 4
All students [TOTAL] = All students
By jurisdiction, 1992, 1994, 1998, 2000, 2002, 2003 and 2005



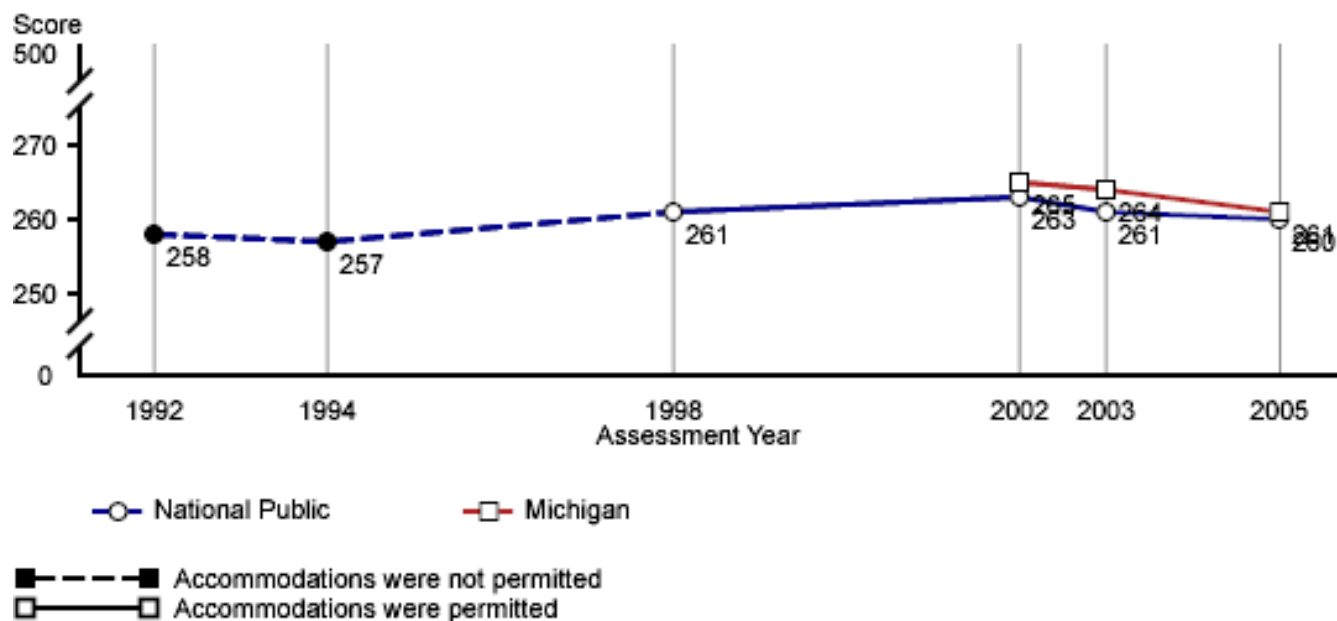
NOTE: The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, 2000, 2002, 2003 and 2005 Reading Assessments.

NAEP Grade 8 Reading Achievement



Average scale scores for reading, grade 8
All students [TOTAL] = All students
By jurisdiction, 1992, 1994, 1998, 2002, 2003 and 2005

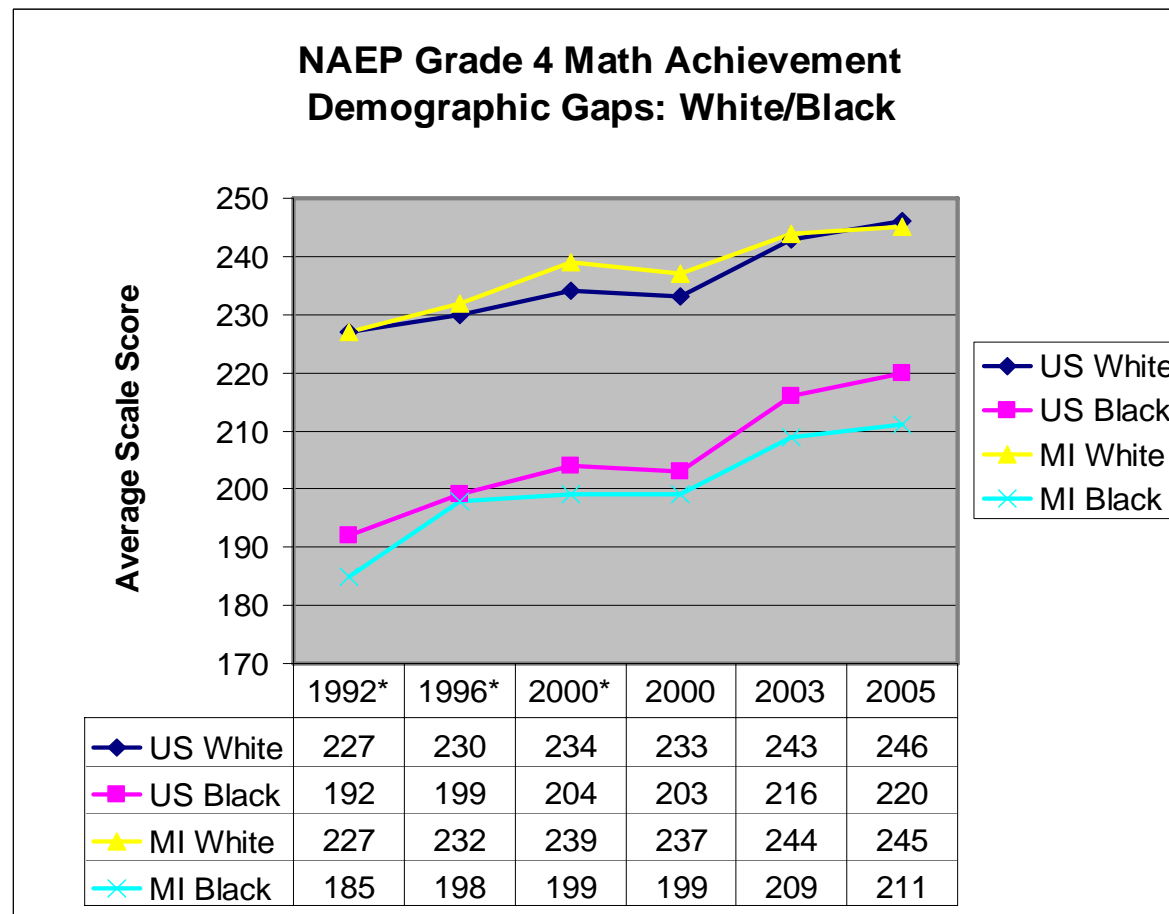


NOTE: The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, 2002, 2003 and 2005 Reading Assessments.



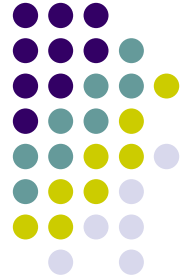
Demographic Gap Analysis G4M: Ethnicity



•Accommodations were not permitted for this assessment.

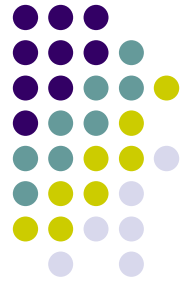
•NOTE: The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.

•SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990, 1992, 1996, 2000, 2003 and 2005 Mathematics Assessments.



Online MEAP

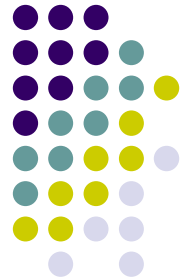
- Describe the Online MEAP Pilot
- Student Performance Comparison
- Results
 - Effects on Student Scoring
 - Surveys and Focus Group Results
- Where do we go from here?



Online MEAP

- Used 19 Freedom to Learn Schools
- One to one computing, HP Laptop - Wireless
- Students and teachers and tech support have experience
- Tested use of wireless laptop assessment environment - surprise!
- 19 MI schools over 2403 students in study
- Tested in 1204 6th Grade ELA and 1163 SS
 - Multiple Choice and Constructed Response
 - Included Automated Essay Scoring

Pearson Educational Measurement



- U.S. Map
- Regional Map
- State Homepage

Michigan

Welcome to the Michigan Educational Assessment Program (MEAP) homepage. Click on the links below to access the services that are currently available to you. Some links may require a user login to access.

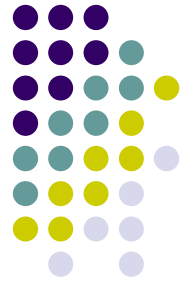
[Practice Center](#)

Enter the MEAP Practice Center which allows you to practice using the eMeasurement Services system without impacting live data.

[Michigan Educational Assessment](#)

Enter the Michigan Educational Assessment testing program.

Pearson Educational Measurement




PEM has:

- Several years experience in state wide online assessment
- TestNav Software in 7 states
 - First use of Intelligent Essay Scoring
- Contractor for MEAP testing
 - Good coordination with P & P

Pearson Educational Measurement









Michigan Educational Assessment Program (MEAP)

- MAIN
- SERVICES
- CUSTOMERS
- FAQ
- SPOTLIGHT
- DEMO
- NEWS
- CONTACT US

- U.S. Map
- Regional Map
- State Homepage

Michigan

Welcome to the Michigan Educational Assessment Program (MEAP) homepage. Click on the links below to access the services that are currently available to you. Some links may require a user login to access.

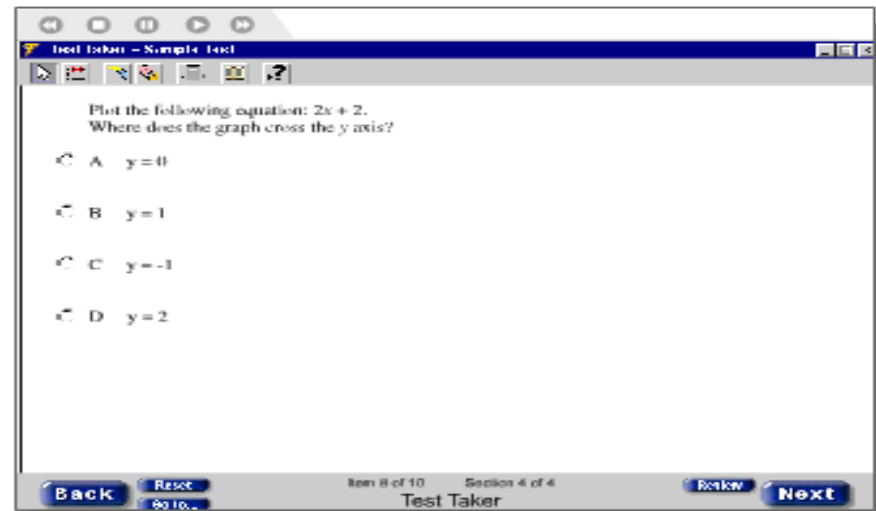
Student/Test Management		System Management	
			
Session Management	Test Delivery	Resources	Security
Assign students to take a test, start and stop the test sessions.	Download TestNav, student tutorials for taking tests on TestNav, take a test.	Access informational resources, user manuals, and documentation.	Manage user profiles, assign roles and data access privileges, reset passwords.

Pearson Educational Measurement



Before starting the TestNav tutorial, let's review how it works.

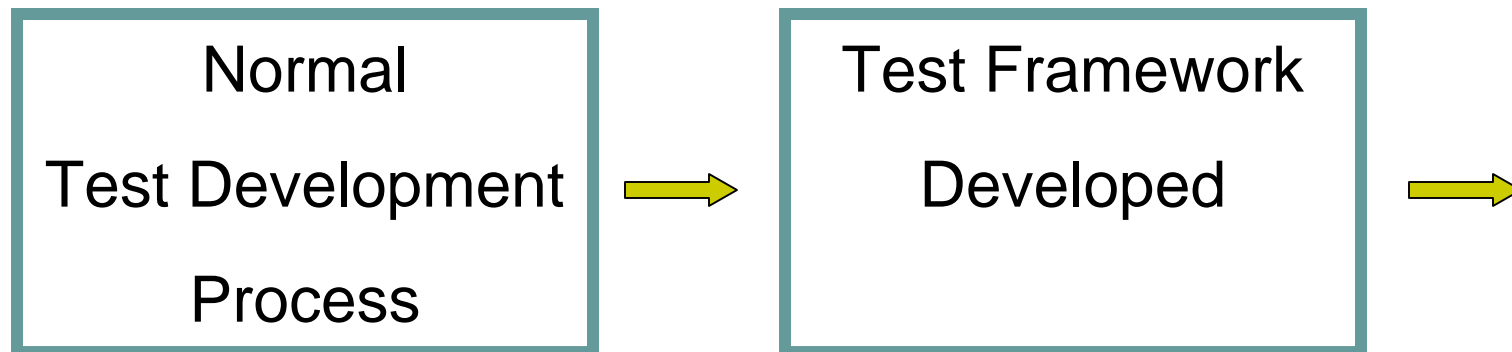
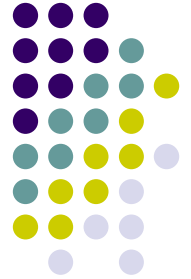
Although, this tutorial is self-paced, you can move through the sections at your own speed.



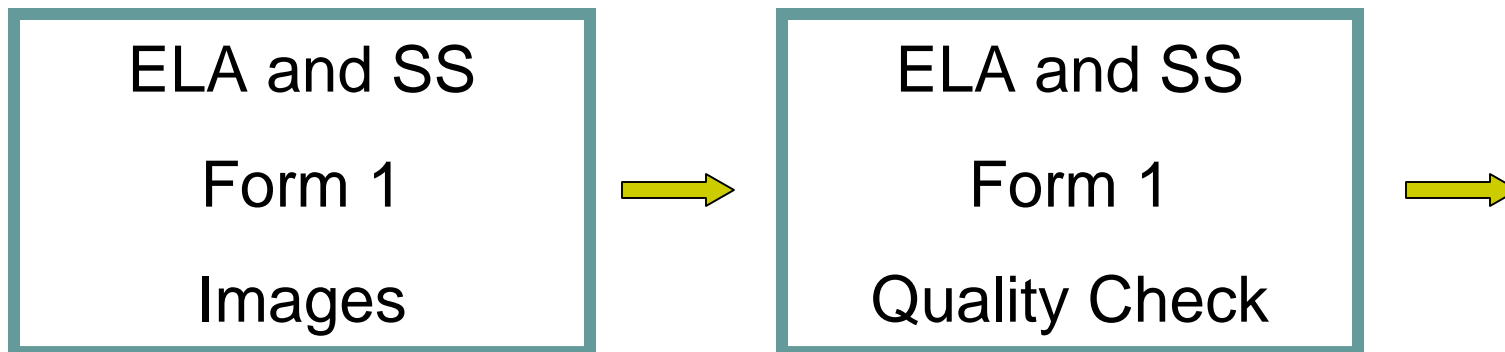
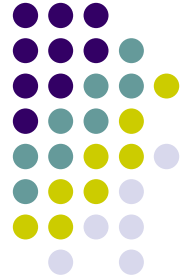
◀ Back

Next ▶

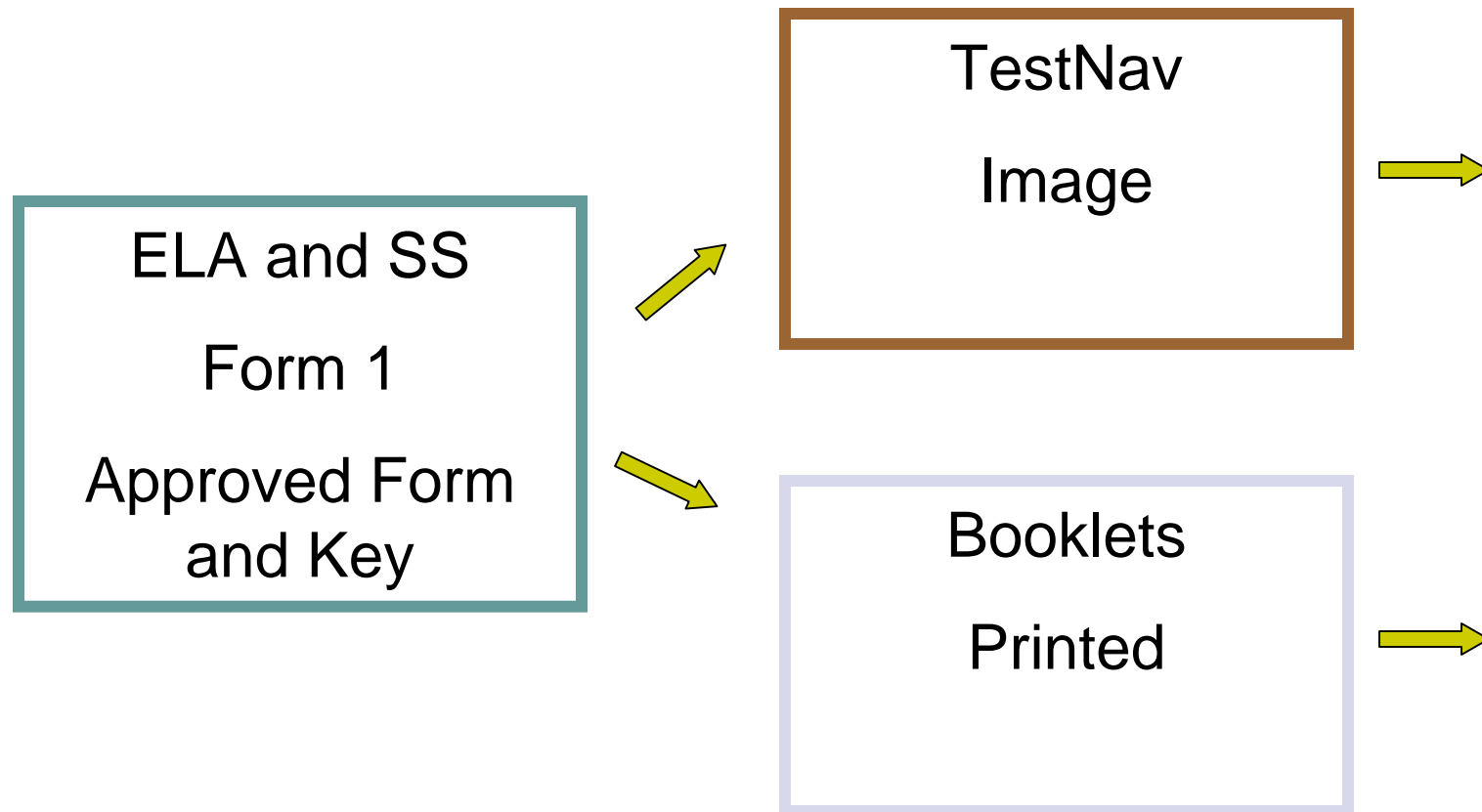
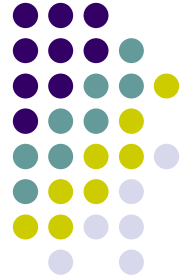
Test Process Slide 1



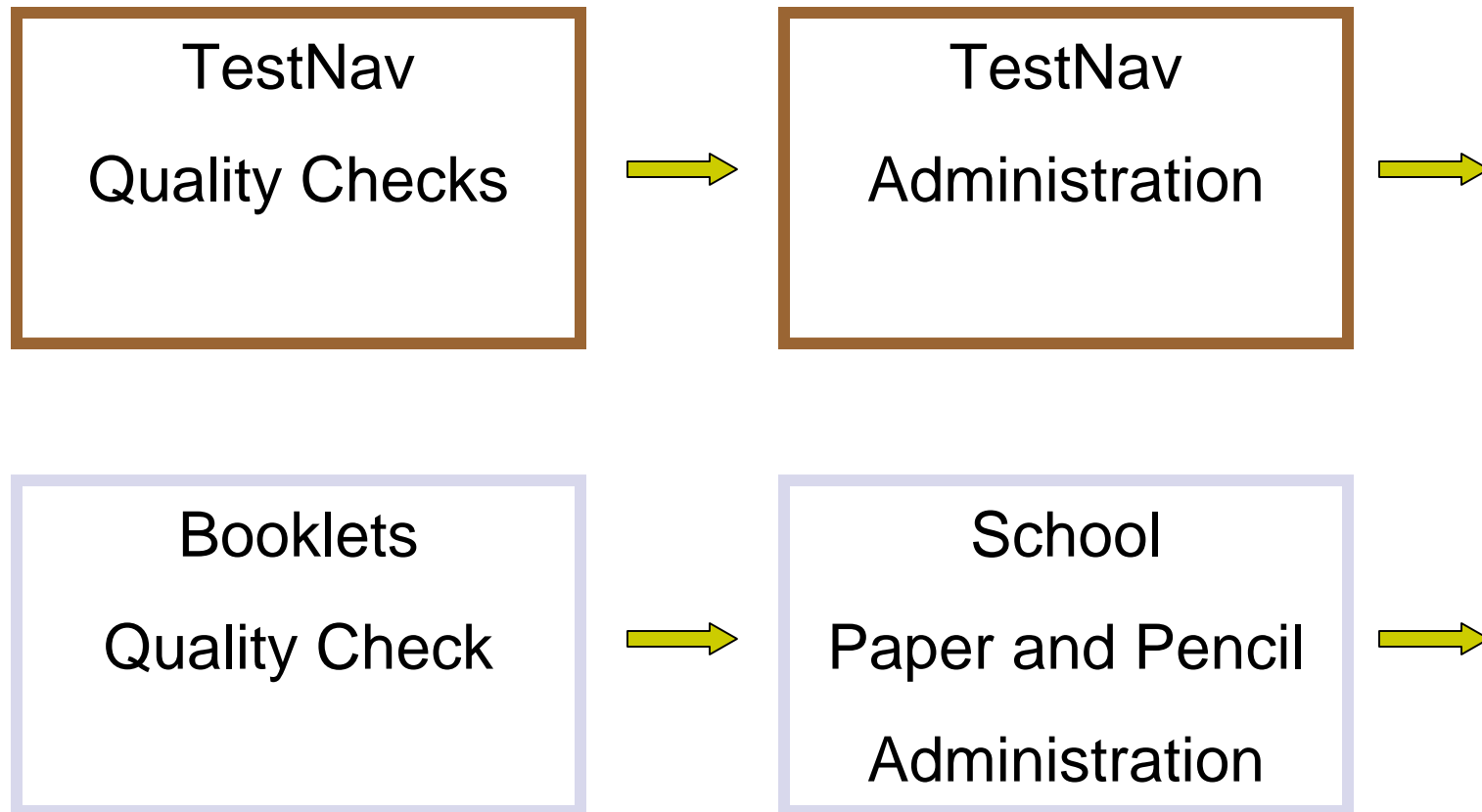
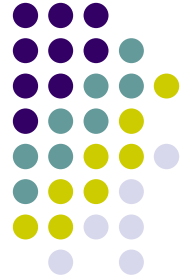
Test Process Slide 2



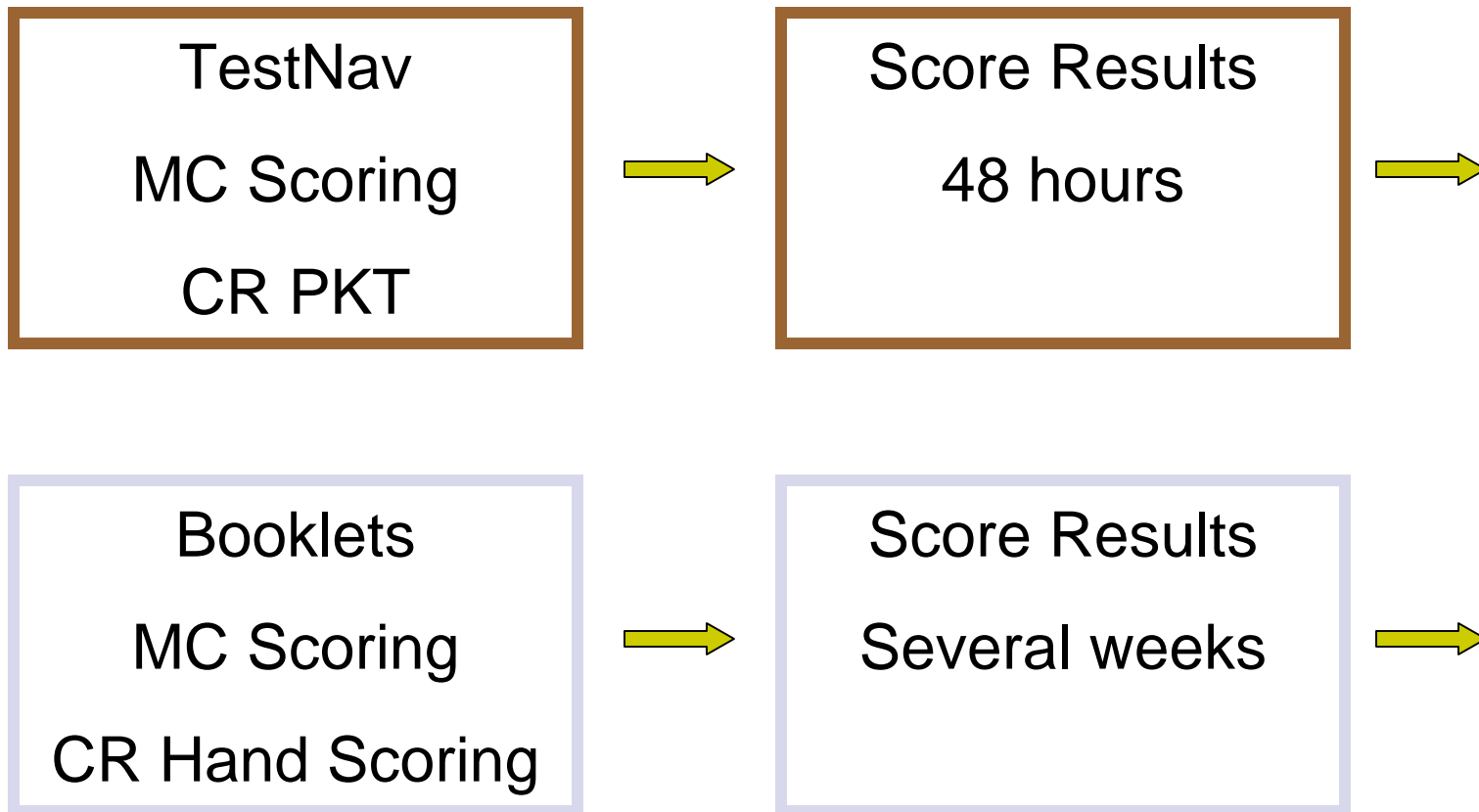
Test Process Slide 3



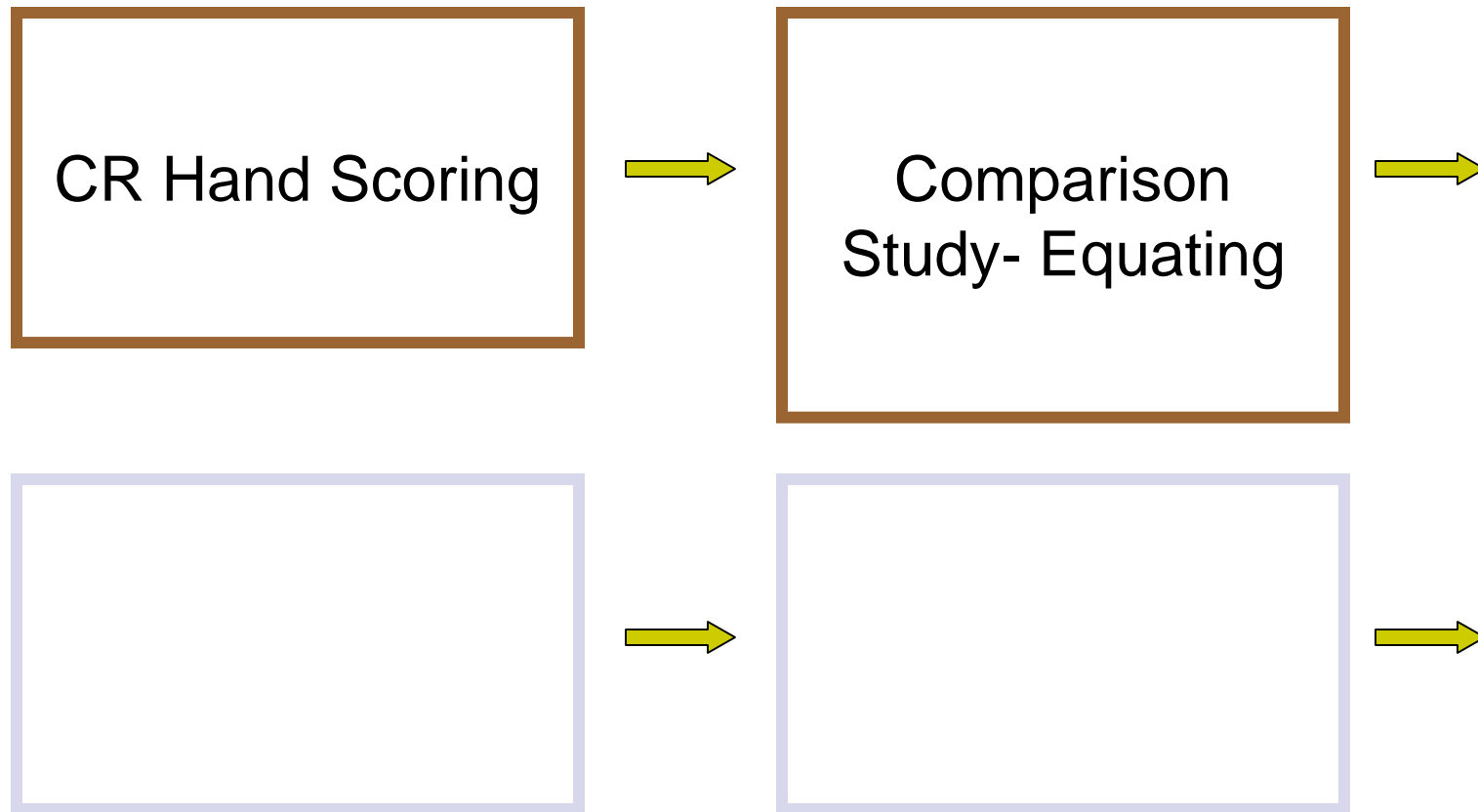
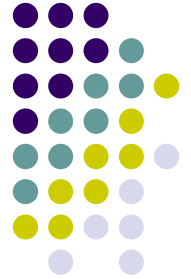
Test Process Slide 4



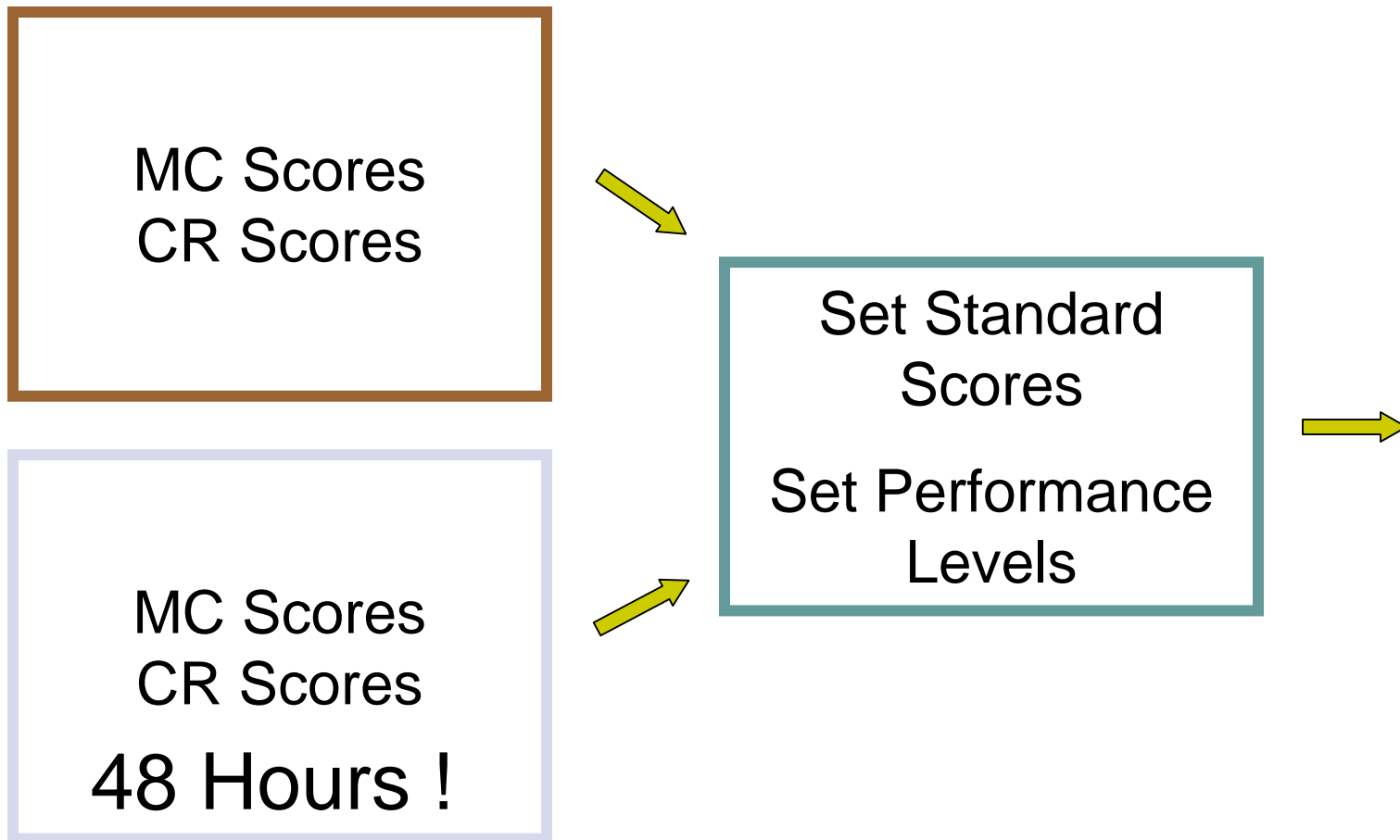
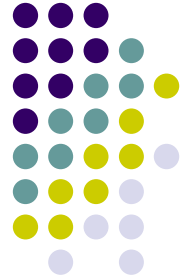
Test Process Slide 5



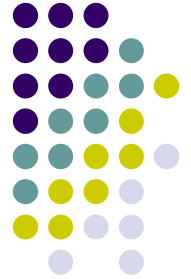
Test Process Slide 6



Test Process Slide 7

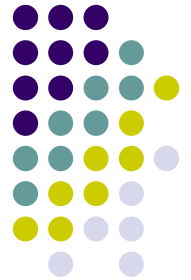



Test Process Slide 8



State wide
Reports Release

Pearson Educational Measurement









Michigan Educational Assessment Program (MEAP)

- MAIN
- SERVICES
- CUSTOMERS
- FAQ
- SPOTLIGHT
- DEMO
- NEWS
- CONTACT US

- U.S. Map
- Regional Map
- State Homepage

Michigan

Welcome to the Michigan Educational Assessment Program (MEAP) homepage. Click on the links below to access the services that are currently available to you. Some links may require a user login to access.

Student/Test Management		System Management	
			
Session Management	Test Delivery	Resources	Security
Assign students to take a test, start and stop the test sessions.	Download TestNav, student tutorials for taking tests on TestNav, take a test.	Access informational resources, user manuals, and documentation.	Manage user profiles, assign roles and data access privileges, reset passwords.

Pearson Educational Measurement



Resources

Click on the links below to access the functions within the Resources system. Accessing these links will require a user login to proceed.

▶ [User's Guide](#)

-Download eMeasurement Services test administration and reporting instructions. (PDF format)

▶ [Infrastructure Guidelines](#)

-Provides information about eMeasurement Services hardware/software requirements, proxy environments, browser setup, and installation of TestNav.

▶ [Proctor Caching User's Guide](#)

-Provides information on the installation and use of eMeasurement Services Proctor Caching software. (PDF format)

▶ [MEAP Online Pilot FAQ](#)

-Provides some quick answers to the most common questions about the MEAP Online Pilot

Click on the links below to access the eMeasurement Services Tutorial

▶ eMeasurement Services Tutorial

View a tutorial of the eMeasurement system.

→ [View Online](#)

↓ [Download \(Mac\)](#)

↓ [Download \(PC\)](#)

Click on the links below to access the functions within the Test Delivery system

▶ TestNav Tutorial - No Sound Version

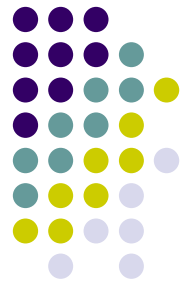
This computer-driven tutorial demonstrates how to use the TestNav delivery system in a text- and graphics-based format. This tutorial can be viewed online or can be downloaded in a Macintosh™ or PC format.

→ [View Online](#)

↓ [Download \(Mac\)](#)

↓ [Download \(PC\)](#)

Pearson Educational Measurement



Creating a New Test Session

Create a New Session

To create a new session, complete the information below. When you are finished, click the Next button to continue.

Test Administration:

School:

Subject:

Test to be Administered:

Session Name:


Example: Smith 2nd Period 8th Grade Math Test

Scheduled Start Date:


Scheduled Start Time:

Estimated Duration:

Location/Room (Optional):



November 2005						
S	M	T	W	T	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			



Pearson Educational Measurement



Adding Students to a Session

Add/Remove Students

Test Administration: Fall 2005

School: School name starts with:

Subject:

Test to be Administered: MEAP ELA Grade 6

Session Name: Backin 2nd Period ELA

Add students to this session and click Next to continue.

Available Students	Students Assigned
Fast Find <input type="text"/>	Student Count: 2
<div><div>Carlo, Frank Y</div><div>Carsosa, Jorge</div><div>Huang, Panakpa</div><div>Krezer, Alana</div><div>Nailor, Benita M</div></div> <div><input type="button" value="Add >"/> <input type="button" value="Add All >>"/> <input type="button" value="Remove <"/></div>	<div>Blyvin, Tania Selman, Ty</div>

[<Previous Step](#)

Pearson Educational Measurement



Session Roster

Session Roster

Test to be Administered: MEAP ELA Grade 6

Session Name: Backin 2nd Period ELA List Get Authorizations

Session created for: OAK PARK SCHOOL DISTRICT/HELM CITY PUBLIC SCHOOL/ELA

Status	Test Code	Start Date	Start Time	Duration	Location	Form Group Type
Not Started	<input type="text" value="R0G417"/>	11/08/2005	8:30 AM	60	Reading Lab	<input type="text" value="Main"/>

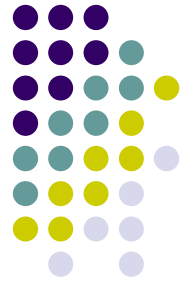
Session Roster

Remove	Resume	Status	Student Name	Unique ID	Login ID	Form or Form Group
<input type="checkbox"/>		Not Started	ELYVIN, TANIA	123-456788058	123-456788058	Main
<input type="checkbox"/>		Not Started	CANTO, FRANK Y	123-486573984	123-486573984	Main
<input type="checkbox"/>		Not Started	CONSECO, JORGE	123-769871234	123-769871234	Main
<input type="checkbox"/>		Not Started	HUANG, PENELOPE C	123-592839485	123-592839485	Main
<input type="checkbox"/>		Not Started	KEEZER, ALANA	123-679999938	123-679999938	Main
<input type="checkbox"/>		Not Started	NAILOR, BENITA M	123-658951295	123-658951295	Main
<input type="checkbox"/>		Not Started	SELLMAN, TY	123-238958734	123-238958734	Main

Restore Apply Add Students Move Students Delete Session Print Start

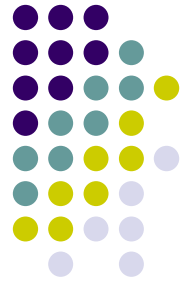
Proctor Caching

Educational Staff Survey and Focus Group



- Surveymonkey.com
- 28/44 School staff responded (63.6%)
 - School Administrator - 7%
 - Test Admin/Proctor - 64%
 - MEAP coordinator - 7%
 - Tech Support - 11%
 - Other - 11%
- Conducted Focus Group Faculty (via videoconference w assistance Sam LaPresto-MASA)

Educational Staff Input - Test Administration



- Survey: Overall test administration and set-up was:
61% {
 - Easy 7%
 - Required some work, but not overly taxing - 54%
 - Too difficult - 18%
 - Wireless networking issues
 - Needed more specific instructions specific to etesting
- Focus Group: some students seemed to hurry, some students did not use scrolling on text box, filled up visible, submit button problem

Pearson Educational Measurement



Test Taker – Sample Test

GENERAL DIRECTIONS

1. Follow the directions on the screen as your teacher reads them.
2. If you cannot see all the instructions or questions at one time, scroll to the right hand side of the window. You can view the entire content and scroll down using the scroll bar. Either click and hold on the up and down arrows of the scroll bar, or click and drag on the scroll bar.
3. The test questions will be shown one at a time, followed by a list of possible answers. You will select only the one answer that you think is the correct answer.

Directions may be displayed at the beginning of the test. During an actual test, follow along as your teacher reads the directions aloud.

[Back](#) [Next](#)

When you click on an answer, it means that additional information is available. To change your answer, click on the new answer. To remove a previously selected answer, click the Reset button. To move to the next question, click the Next button at the bottom of the screen. To return to the previous question, the Previous button, if it is present, will allow you to do so.

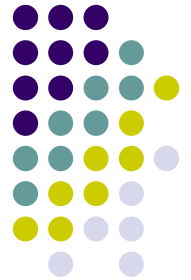
Mark for Review. At the end of the section, you will be able to review this item.

By closing the window, switching applications, or using the task bar, you will be able to do so.

GENERAL DIRECTIONS

[Next](#)

Pearson Educational Measurement



Test Taker – Sample Test

Evaluate the expression: $a + bc^2$ when $a = 6$, $b = 4$, and $c = 2$ } Question

☐ A 15
☐ B 22
☐ C 17
☐ D 26 } Choices

Each question is displayed one at a time at the top of the screen. The multiple-choice options are shown with the question.

◀ Back Next ▶

Back Reset Go to Item 1 of 10 Section 4 of 4 Review Next

Pearson Educational Measurement



Test Taker – Sample Test

White Fang
by
Jack London

Dark spruce forest frowned on either side of the frozen waterway. The trees had been stripped by a recent wind of their white covering of frost, and they seemed to lean toward each other, black and ominous, in the fading light. A vast silence reigned over the land. The land itself was a desolation, lifeless, without movement, so lone and cold that the spirit of it was not even that of sadness. There was hint in it of laughter, but of a laughter more terrible than any sadness – a laughter that was mirthless as the

What is the meaning of the word *prominent*?

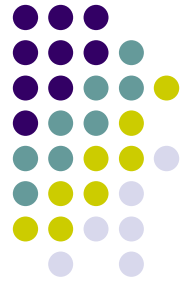
- ☐ A most frightening
- ☐ B most noticeable
- ☐ C most incredible
- ☐ D most useful

For some questions, TestNav displays a split screen. A reading passage or other information is shown in the top screen, and the accompanying question in the bottom screen.

Back Next

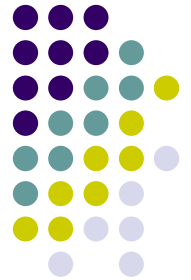
Back Reset Go to Item 1 of 10 Section 3 of 4 Review Next

Test Taker



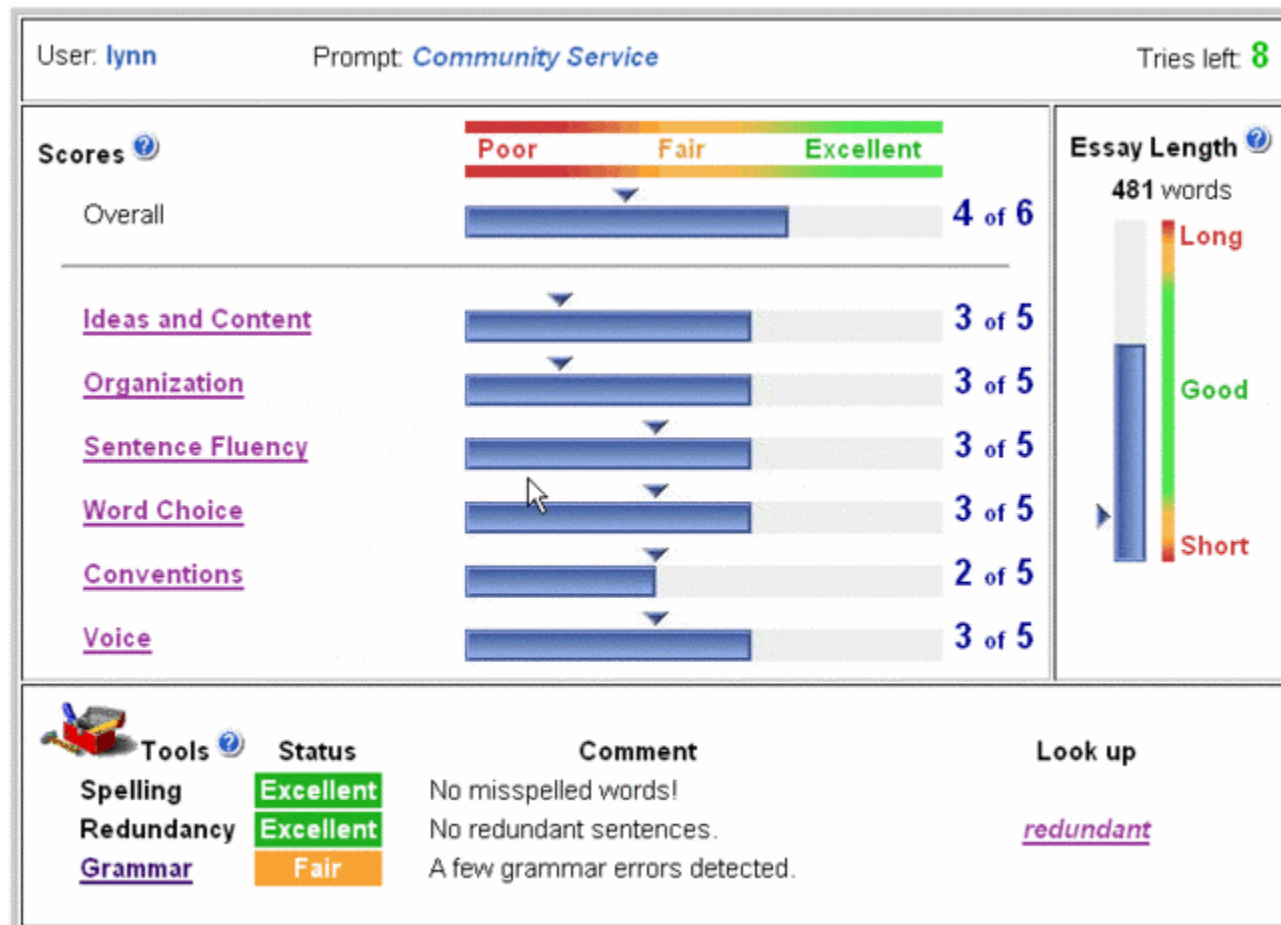
Intelligent Essay Assessor

- Features of the Intelligent Essay Assessor™
- Proven reliability
- Individually tailored feedback is returned in seconds
- A hosted application capable of scoring millions of essays a day
- Built in detectors for: plagiarism, abnormal English, highly unusual essays
- Able to provide practice and drill with a textbook and study guide – not used in MEAP scoring

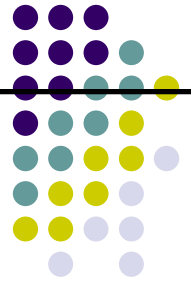


Intelligent Essay Assessor

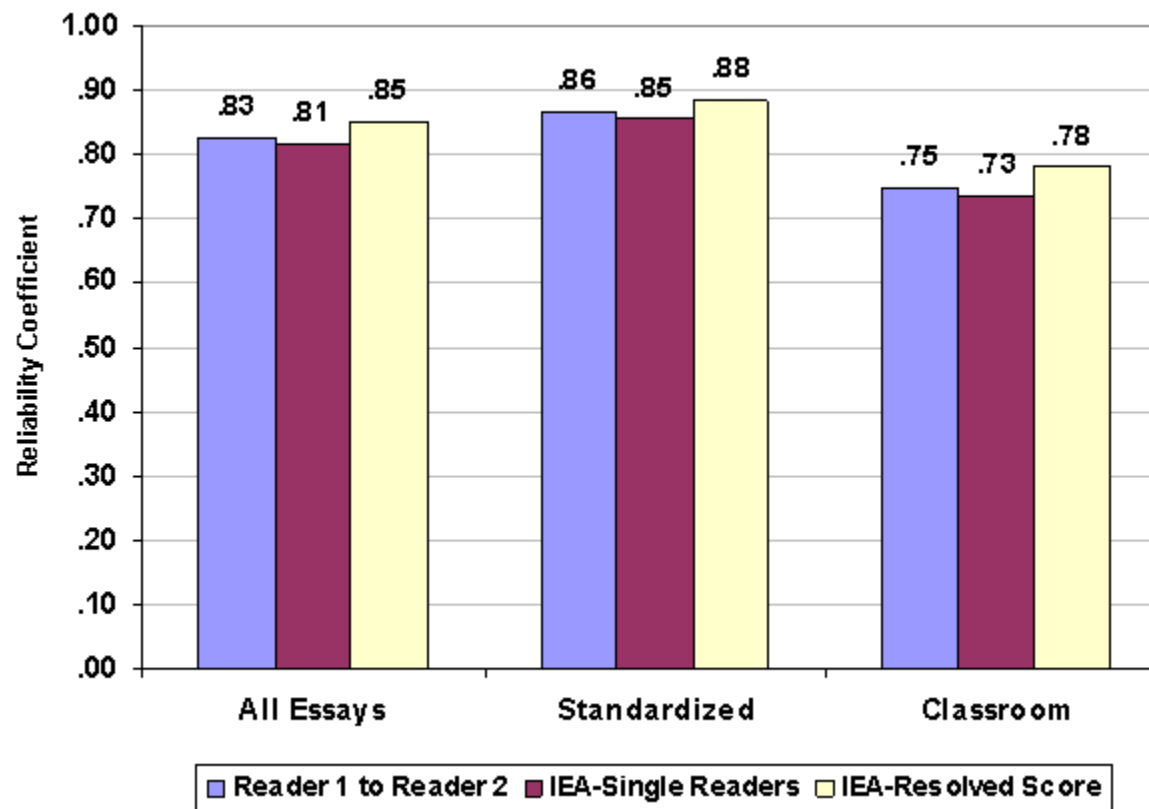
Analytic or Holistic Scoring



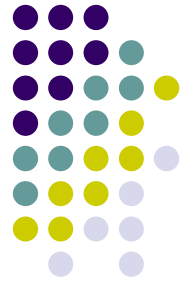
Intelligent Essay Assessor



Inter-rater reliability for resolved reader scores

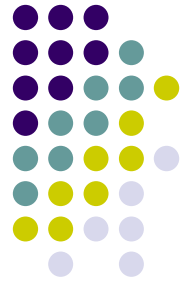


Overall Comparison Online/Paper (all 6th Graders Form 1)



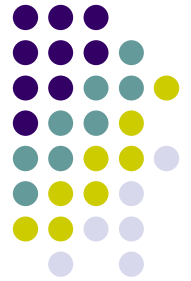
Subject	Mean Online Raw Score (N)	Mean Paper Raw Score (N)	Difference Paper- Online
Social Studies	28.96 (1095)	29.60 (11272)	.64
Reading	26.12 (1133)	26.47 (42872)	.35
Writing	6.72 (1133)	7.11 (42872)	.39

Overall Comparability Methodology



- PEM conducted comparability study
 - “matched group” design using 5th Grade Social Studies Test from previous year (same cohort, same test), matched on demographics
 - Bootstrap methodology of Dorans and Lawrence (1990) with 100 replications of matched samples - analyses was conducted using the difference between equating function and the identity transformation, and divide this difference by the standard error of equating (with $\pm 2se$)

Overall Comparability Results



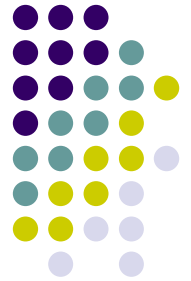
- Scores were sufficiently comparable between online and paper scoring for the same conversion tables to be used
- Especially since MEAP requested that PEM use the higher of the two scores on CR
- The online test was very slightly more difficult with about a .5 raw score difference.



Comparability Discussion

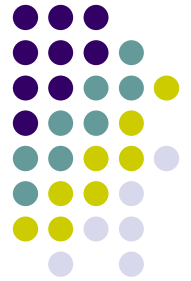
- Why the half point difference?
 - Simply a new testing environment for many students
 - Familiarity with online assessment
 - “Rushing” reported by educational staff
 - Need to warn students not to rush
 - Unfamiliarity with split screen for reading
 - Some schools did make good use of paper for notes and drafts - (administrative errors due to lack of manual and training)

Constructed Response Results



- Analyses of CR items in reading, writing and ss were scored lower by IEA (higher of hand v. computer score was assigned so no penalty to participating students)
- Within the 48 hour period, agreements were:
 - SS Exact 60%, Adjacent 99.5% (lower)
 - Reading CR1 Exact 46%, Adjacent 94%
 - Writing CR1 Exact 65%, Adjacent 99%
 - Writing CR2 Exact 48%, Adjacent 97%

Constructed Response Results

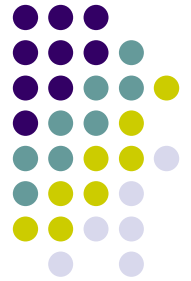


- Why poor performance of IEA?
 - Suspected that we had inadequate field test sample
 - (especially inadequate number of responses in extreme ranges (1,6))
 - PEM did an additional study using n=200 and n=500 sample responses to train the IEA engine
 - Results

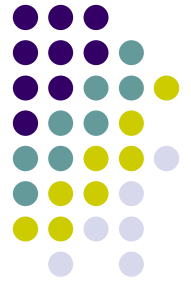
● Subject	200 IEA Exact	500 IEA Exact	Human Exact
● SS	61	67	71
● Reading	51	56	59
● Writing 1	71	75	65
● Writing 2	62	73	75

Adjacent were all over 98% (except for Reading which was 89%)

Constructed Response Results

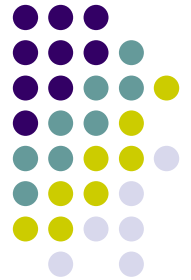


- Lessons learned
 - From focus group data and student survey data -
 - Some students are not probably not ready for CR via TestNav
 - Students did not use draft writing paper (failure to follow directions)
 - Students did not have adequate practice writing CR using TestNav (ie scrolling)
 - To properly train the engine for MEAP CR requires a minimum $n \geq 500$ in most cases
 - Two questions to answer:
 - 1. For those students ready, should a) they be allowed to take the CR via TestNav? b) should IEA scoring within 48 hours be allowed?
 - 2. Should we also allow a paper writing and scoring option for students not ready (is grade 6 too early for TestNav essay scoring and input?)



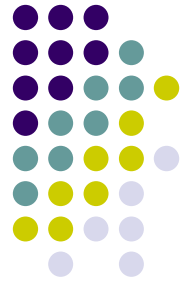
Overall Evaluation of Pilot

- Student and Faculty opinions of online testing were overwhelmingly positive (90% of students would recommend or strongly recommend, 93% of faculty would recommend or recommend with some reservations, less than 4% would strongly not recommend)
- The opinion of faculty was that 39% finished the exams in less time than p&p, 29% about the same time, and 21% thought online testing took more time
- For a first time, a surprising number of students took advantage of unique helping tools
 - For example, SS, Highlighter 40%, Eraser 31%, Answer Elimination 44%, those that reported using no tools: 36%



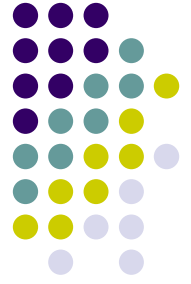
Overall Evaluation of Pilot

- Some students do not appear to be ready for online essay writing at grade 6:
 - Responding to the question: It is easier for me to write an essay on paper than on the computer yielded
Strongly Agree Agree Neutral Disagree Strongly Disagree
22 14 26 38 0
(36% prefer paper??)



Summary

- Student scores between online and p&p are sufficiently comparable
- CR IEA scoring needs more work and considerations:
 - Students need better instructions, use of drafts
 - Students need more experience with TestNav
 - Some students may not have pre-requisite skills
 - Field trial samples need to be 500 or greater



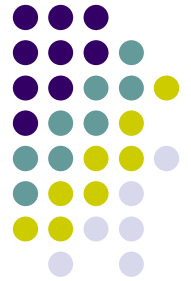
A few surprises

- We went “wireless”
- We went cordless
- Some students went too fast in the eyes of the proctors
- Some students didn’t know they could scroll down and add more text to their CR responses



Lessons Learned

- To PEM: Need an online essay screening proficiency test
- For schools/OEAA:
 - Need to start qualifying equipment/network configurations sooner
 - Need to make sure students do practice test that includes a constructed response test
 - Need to change configurations to match Test Software - screen savers off, firewall changes possible, other security issues, may need to enlist students in these changes
 - Need closely coordinated tech support



Where do we go from here?

- Evaluation shows that online testing method appears valid
- We've learned a lot about some of the problems with online testing thanks to the FTL pioneers
- If we were to offer this in the future, what is the best method to phase this in?
 - Special ed accommodations? ELL, ELPA, pick a grade, subject?